Deep Uncertainty Quantification: With an Application to Integrated Assessment Models

University of Geneva

Simon Scheidegger
Department of Economics, University of Lausanne

joint work with:
Aleksandra Friedl
ifo Institute

Felix Kübler
Institute for Banking and Finance, University of Zurich; and the SFI

Takafumi Usui
Institute for Banking and Finance, University of Zurich

March 26th, 2025

Climate Change



- Global warming is a growing threat (to economic well being,...).
- Portends massive human suffering.
- Tipping points place us in catastrophic danger.
- It will impact current and future generations in different regions very differently.
- A recent review by Fernandez-Villaverde, Gillingham, Scheidegger (2024): `https://www.nber.org/papers/w32963`

Why Uncertainty Quantification in a Stochastic IAM?

- **Carbon taxation**: Widely supported policy response to a climate change problem.
- Level of carbon tax is based on the social cost of carbon (SCC - marginal loss caused by an extra ton of CO2 emissions).
- SCC values are computed with **integrated assessment models** (IAMs) that link economy and climate.
- IAMs are subject to **significant parametric uncertainty, model uncertainty, and climate risks of tipping points**.
- **Numerically difficult task** since IAMs are complex non-linear models that are highly susceptible to the curse of dimensionality.

Our contribution

- **Generic global solution method** to efficiently compute global solutions to stochastic IAMs that include economic and climate risks.
- **Perform global sensitivity analysis** with respect to uncertain parameters most discussed in the literature.
    1. To alleviate the curse of dimensionality we exploit a novel numerical approach based on **deep learning**.
    2. To take into account parametric uncertainty, we add the model parameters as **pseudo-states** to the state space $\rightarrow$ we need to solve the model only once.
    3. We construct **surrogate models** for the Social Cost of Carbon (and other quantities of interest) as a function of all relevant model parameters by using **Gaussian Process regression and Bayesian active learning**.

Related Literature

- Literature on IAMs with economic and climate risks:
  - IAMs taking risks into account are focusing only on one type of the risk, such as long-run growth uncertainty (Jensen and Traeger, 2014) or climate tipping points (Lemoine and Traeger, 2016).
  - IAM that includes both economic and climate risks (Cai and Lontzek, 2019) were solved using thousands of node-hours on a supercomputer.[1]
- Literature on parametric uncertainty quantification:
  - Local sensitivity analysis based on Monte-Carlo simulations (Anthoff et al., 2009; Ackerman et al., 2010; Gillingham and Stock, 2018).
  - Global sensitivity analysis available only for deterministic IAMs (Anderson et al., 2014; Butler et al., 2014; Miftakhova, 2021).

→ We introduce a generic way of performing global sensitivity analysis for stochastic IAMs.

[1] One model solution corresponds to about 11 years on a laptop.

Outline of this talk

1. Integrated Assessment Models (IAMs) in a Nutshell
2. Deep Equilibrium Nets for IAMs
3. Global Uncertainty Quantification
4. Some Tentative Results

Integrated Assessment Models (IAMs) in a Nutshell

Integrated Assessment Models (IAMs)

- Pioneered by Nordhaus (DICE, RICE): Quantitative, often numerical.
- Key components:
  - (Simplified) climate system.
  - (Stylized) economic model of emissions AND damages.
- Economic model: needs to be dynamic, forward-looking, possibly allowing stochasticity (temperature variations, disasters, TFP), and potentially heterogeneous agents ("who is how impacted?").

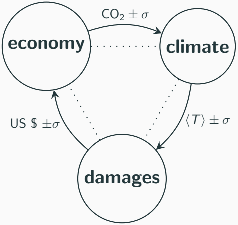
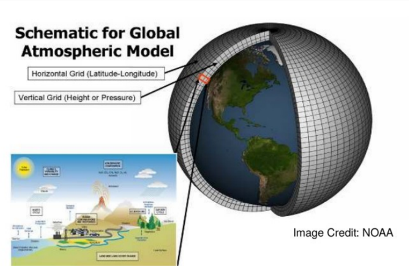


Figure 1: Stylized representation of an IAM.

Global Atmospheric Models



- AOGCM: Atmosphere Ocean Global Climate Models.
- Back-bone of IPCC (Intergov. Panel on Climate Change).
- Costly & complex (few hours wall-clock per model year & $\approx 1mio$ lines of code)
- One 1y of simulation uses about 3,000 CPUh.
- Simplifications needed (DICE, FAIR, ...).

$\rightarrow$ Need a climate model that is numerically affordable!

The blue marble in 5 state variables (I)

- Climate: functional form of DICE-2016 (Nordhaus, 2017a).
- Parameterization by Folini et al. (2024) (match CMIP5 output).
- DICE-2016 models the carbon cycle via three carbon reservoirs.
- Atmosphere (A), Upper ocean (U), Lower ocean (L).
- Aggregate distribution of carbon in the world given by 3 boxes: $M_t = (M_{A,t}, M_{U,t}, M_{L,t})^T$
- Carbon concentrations evolve as linear dynamic system:

$$M_{t+1} = (I + \Delta_t \cdot B) \cdot M_t + \Delta_t \cdot E_t.$$

- Carbon Emissions (econ. activity + exogenous source): $E_t = E_{ind,t} + E_{Land}$.
- Can easily afford to add more reservoirs (Eftekhari et al. (2024), in preparation).

The blue marble in 5 state variables (II)

- The two-layer energy balance model in DICE-2016 reads as

$$T_{t+1}^{\mathrm{AT}} = T_t^{\mathrm{AT}} + \Delta_t \cdot c_1 \left(F_t - \lambda T_t^{\mathrm{AT}} - c_3 \left(T_t^{\mathrm{AT}} - T_t^{\mathrm{OC}} \right) \right),$$
$$T_{t+1}^{\mathrm{OC}} = T_t^{\mathrm{OC}} + \Delta_t \cdot c_4 \left(T_t^{\mathrm{AT}} - T_t^{\mathrm{OC}} \right).$$

- Radiative forcing:

$$F_t = F_{\mathrm{2XCO2}} \frac{\log(M_t^{\mathrm{AT}} / M_{\mathrm{EQ}}^{\mathrm{AT}})}{\log(2)} + F_t^{\mathrm{EX}}.$$

CDICE vs. DICE-2016

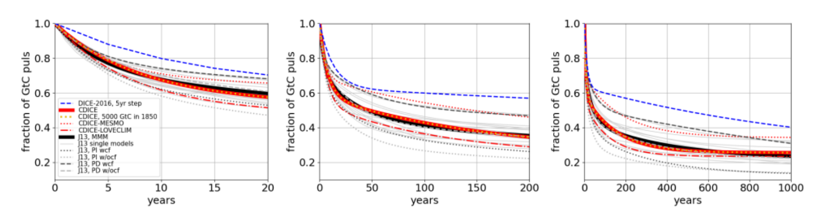


Figure 2: A comparison of the calibrated climate emulator CDICE versus DICE-2016 (Folini et al., 2024).

A Source of Uncertainty: Equilibrium Climate Sensitivity

IPCC AR5 gives a 'likely range' of [1.5K, 4.5K]

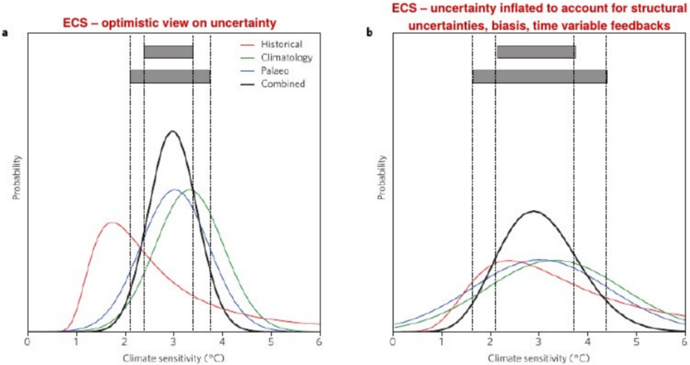
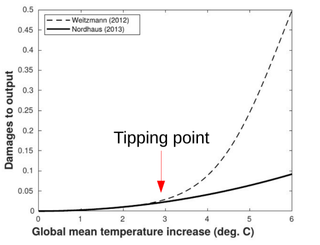


Figure 5 | Illustrative example of combining multiple constraints for climate sensitivity. Overall PDFs are the products of three PDFs based on the historical warming, climatological constraints on mostly GCMs, and palaeoclimate. Grey ranges at the top indicate a 'likely' (66%) and 'very likely' (90%) combined range. **a**,**b**, Constraint based on an optimistic interpretation of uncertainty ranges and assuming full independence (**a**) and on inflated ranges (**b**) to account for structural uncertainties, and with historical estimates inflated and scaled up to account for observation biases, and feedbacks varying from historical to future and across forcings. See Methods for details.

Knutti et al. 2017, Nature, doi: 10.1038/NGEO3017

Another Source of Uncertainty: Damage Functions

See also Tol (2009), Hänsel et al. (2020)



- Weitzmann (2012):

$$\Omega_t\left(T_{\mathrm{AT},t}\right) = \frac{1}{1 + \left(\frac{1}{\psi_1} T_{\mathrm{AT},t}\right)^2 + \left(\frac{1}{2 \cdot TP_t} T_{\mathrm{AT},t}\right)^{6.754}}.$$

- Nordhaus (2013):

$$\Omega_t^N\left(T_{\mathrm{AT},t}\right) = \frac{1}{1 + \pi_1 T_{\mathrm{AT},t} + \pi_2 T_{\mathrm{AT},t}^2}.$$

- Degree of convexity is of key importance in determining the optimal taxes, not only the level.
- Tipping points: include losing much of the Amazon rain forest, faster onset of El Nino, the reversal of the Gulf Stream, etc.

Quantifying uncertainty about ECS

- Temperature equation:

$$T_{\mathrm{AT},t+1} = \left(1 - \varphi_{21} - \varphi_1 \frac{F_{\mathrm{2xco2}}}{\Delta T_{\mathrm{AT},\times 2}}\right) T_{\mathrm{AT},t} + \varphi_{21} T_{\mathrm{OC},t} + \tag{1}$$

$$\varphi_1 \left(F_{\mathrm{2xco2}} \log_2 \left(\frac{M_{\mathrm{AT},t}}{M^*_{\mathrm{AT}}}\right) + F_{\mathrm{EX},t}\right) \tag{2}$$

- Follow Roe and Baker (2007) to make ECS stochastic.
- $\mathrm{ECS} = \frac{\lambda_0}{1-f} F_{\mathrm{2XCO2}}$, with climate feedback parameter $f \sim \mathcal{N}(\mu_f, S_f)$.
- This is controversial, see Zaliapin and Ghil (2010) and Roe and Baker (2011) and Zaliapin and Ghil (2011).

Bayesian learning mechanics

- Temperature evolves as:

$$T_{\mathrm{AT},t+1} = \left(\varphi_1 \frac{F_{\mathrm{2xco2}}}{\Delta T^0_{\mathrm{AT},\times 2}} \tilde{f}_{t+1} + 1 - \varphi_{21} - \varphi_1 \frac{F_{\mathrm{2xco2}}}{\Delta T^0_{\mathrm{AT},\times 2}}\right) T_{\mathrm{AT},t}$$
$$+ \varphi_{21} T_{\mathrm{OC},t} + \varphi_1 \left(F_{\mathrm{2xco2}} \log_2 \left(\frac{M_{\mathrm{AT},t}}{M^*_{\mathrm{AT}}}\right) + F_{\mathrm{EX},t}\right) + \tilde{\epsilon}_{T,t+1}$$

  where $\tilde{f}_{t+1} \sim \mathcal{N}\left(\mu_{f,t}, S_{f,t}, \underline{f}, \overline{f}\right)$ - uncertain climate sensitivity, $\tilde{\epsilon}_{T,t+1}$ - shock to the temperature.
- The agent observes a realisation of the shocks in climate sensitivity and temperature at the beginning of the period t before the decisions are made.
- Under Gaussian assumption analytic formulas for updating:

$$\mu_{f,t+1} = \frac{S_{\epsilon_T}\mu_{f,t} + \varphi_{1C} T_{\mathrm{AT},t}\left(\varphi_{1C} T_{\mathrm{AT},t}\tilde{f}_{t+1} + \tilde{\epsilon}_{T,t+1}\right) S_{f,t}}{S_{\epsilon_T} + \left(\varphi_{1C} T_{\mathrm{AT},t}\right)^2 S_{f,t}}$$
$$S_{f,t+1} = \frac{S_{\epsilon_T} S_{f,t}}{S_{\epsilon_T} + \left(\varphi_{1C} T_{\mathrm{AT},t}\right)^2 S_{f,t}}$$

A Stochastic IAM with Bayesian Learning about the ECS

$$V\left(\boldsymbol{X}_t\right)^{1-1/\psi} = \max_{C_t, K_{t+1}, \mu_t} \left\{ \left(\frac{C_t}{L_t}\right)^{1-1/\psi} L_t + e^{-\rho}\mathbb{E}_t\left[V\left(\boldsymbol{X}_{t+1}\right)^{1-\gamma}\right]^{\frac{1-1/\psi}{1-\gamma}} \right\}$$

s.t.

$$\left(1-\Theta\left(\mu_t\right)\right)\Omega_t\left(T_{\mathrm{AT},t}\right)K_t^{\alpha}\left(A_t L_t\right)^{1-\alpha} - C_t - I_t = 0$$

$$\left(1-\delta\right)K_t + I_t - K_{t+1} = 0$$

$$1-\mu_t \geq 0$$

$$\left(1-\phi_{12}\right)M_{\mathrm{AT},t} + \phi_{21}M_{\mathrm{UO},t} + \left(1-\mu_t\right)\sigma_t K_t^{\alpha}\left(A_t L_t\right)^{1-\alpha} + E_{\mathrm{Land},t} - M_{\mathrm{AT},t+1} = 0$$

$$\phi_{12}M_{\mathrm{AT},t} + \left(1-\phi_{21}-\phi_{23}\right)M_{\mathrm{UO},t} + \phi_{32}M_{\mathrm{LO},t} - M_{\mathrm{UO},t+1} = 0$$

$$\phi_{23}M_{\mathrm{UO},t} + \left(1-\phi_{32}\right)M_{\mathrm{LO},t} - M_{\mathrm{LO},t+1} = 0$$

$$\left(1-\varphi_{21}-\varphi_{1C}\right)T_{\mathrm{AT},t} + \varphi_{21}T_{\mathrm{OC},t} + \varphi_1\left(F_{\mathrm{2xco2}}\log_2\left(\frac{M_{\mathrm{AT},t}}{M_{\mathrm{AT}}^*}\right) + F_{\mathrm{EX},t}\right) + \varphi_{1C}\tilde{f}_{t+1}T_{\mathrm{AT},t} + \tilde{\epsilon}_{T,t+1} - T_{\mathrm{AT},t+1} = 0$$

$$\varphi_4 T_{\mathrm{AT},t} + \left(1-\varphi_4\right)T_{\mathrm{OC},t} - T_{\mathrm{OC},t+1} = 0$$

$$\frac{S_{\epsilon_T}\mu_{f,t} + \varphi_{1C}T_{\mathrm{AT},t}\left(\varphi_{1C}T_{\mathrm{AT},t}\tilde{f}_{t+1} + \tilde{\epsilon}_{T,t+1}\right)S_{f,t}}{S_{\epsilon_T} + \left(\varphi_{1C}T_{\mathrm{AT},t}\right)^2 S_{f,t}} - \mu_{f,t+1} = 0$$

$$\frac{S_{\epsilon^T}S_{f,t}}{S_{\epsilon^T} + \left(\varphi_{1C}T_{\mathrm{AT},t}\right)^2 S_{f,t}} - S_{f,t+1} = 0$$

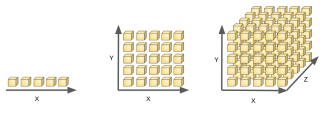
$$\tilde{f}_{t+1} \sim \mathcal{N}\left(\mu_{f,t}, S_{f,t}, \underline{f}, \overline{f}\right), \tilde{\epsilon}_{T,t+1} \sim \mathcal{N}\left(0, S_{\epsilon^T}\right)$$

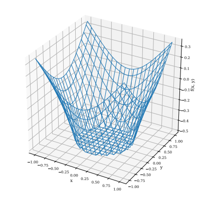
- $X_t = \left[k_t, M_{\mathrm{AT},t}, M_{\mathrm{UO},t}, M_{\mathrm{LO},t}, T_{\mathrm{AT},t}, T_{\mathrm{OC}}, \mu_{f,t}, S_{f,t}, t; \theta\right]$.
- Minimal model has a **9+N-dimensional state space**.

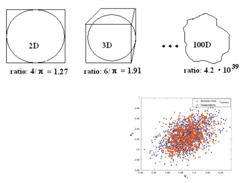
Deep Equilibrium Nets for IAMs

Roadblocks for Performing UQ in Stochastic IAMs

1. Models suffer from the curse of dimensionality.
2. Models suffer from non-linearities.
3. Have to approximate and interpolate high-dimensional functions on irregular-shaped geometries.







→ "Deep Equilibrium Nets" (Azinovic et al., 2022).

Deep Equilibrium Nets & IAMs

- DEQNs can solve nonlinear stochastic models globally, even if geometry is oddly-shaped, in minutes to hours on a laptop.
- Allows to deal with large state spaces (up to dozens of state variables).
- Add pseudo-states for uncertainty quantification at a small extra cost.
- Only one SINGLE model evaluation needed.
- The remaining UQ tasks are just "post-processing", and come at low computational costs, as we can query a surrogate.

Deep equilibrium nets (Azinovic et al., 2022)

A **functional rational expectations equilibrium**: $\{f_i\}_{i=1}^{N_{\text{out}}}$, where

$$f_i : \mathcal{D} \subset \mathbb{R}^{N_{\text{in}}} \to \mathbb{R} : \underbrace{\mathbf{x}}_{\text{state}} \to \underbrace{f_i(\mathbf{x})}_{\text{endogenous variables}}, \text{ s.t. : } \underbrace{\mathbf{G}(\mathbf{x}, f_1, \ldots, f_{N_{\text{out}}}) = 0}_{\text{equilibrium conditions}}$$

A **deep equilibrium net**: $\mathcal{N}_{\boldsymbol{\rho}}$, where

$$\mathcal{N}_{\boldsymbol{\rho}} : \mathcal{D} \subset \mathbb{R}^{N_{\text{in}}} \to \mathbb{R}^{N_{\text{out}}} : \underbrace{\mathbf{x}}_{\text{state}} \to \underbrace{\mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x})}_{\text{approximate endogenous variables}} \approx \begin{bmatrix} f_1(\mathbf{x}) \\ \ldots \\ f_{N_{\text{out}}}(\mathbf{x}) \end{bmatrix}$$

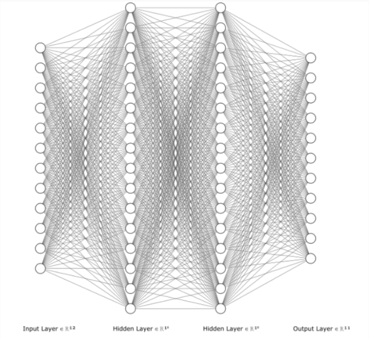
Preview of key ideas:

1. Use the definition of the equilibrium functions, *i.e.* the implied error in the optimality conditions, as loss function.
2. Learn the equilibrium functions with stochastic gradient descent.
3. Take the data points from a simulated path.

What is a deep neural net?

$$
\begin{aligned}
&\textbf{input} := \mathbf{x} \to \phi^1(W^1_{\boldsymbol{\rho}}\mathbf{x} + \mathbf{b}^1_{\boldsymbol{\rho}}) =: \textbf{hidden 1} \\
&\to \textbf{hidden 1} \to \phi^2(W^2_{\boldsymbol{\rho}}(\textbf{hidden 1}) + \mathbf{b}^2_{\boldsymbol{\rho}}) =: \textbf{hidden 2} \\
&\to \textbf{hidden 2} \to \phi^3(W^3_{\boldsymbol{\rho}}(\textbf{hidden 2}) + \mathbf{b}^3_{\boldsymbol{\rho}}) =: \textbf{output}
\end{aligned}
$$

The neural net is then given by the choice of activation functions and the parameters $\boldsymbol{\rho}$.



How to find good parameters ρ

The standard way:

Step 1 : get "labeled data" $\mathcal{D} := \{(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_{|\mathcal{D}|}, \mathbf{y}_{|\mathcal{D}|})\}$ where $\mathbf{y}_i = \mathbf{f}(\mathbf{x}_i)$ is the correct output for input $\mathbf{x}_i$.

Step 2 : Define a loss function, for example:

$$\mathsf{l}_{\boldsymbol{\rho}} := \frac{1}{|\mathcal{D}|} \sum_{(\mathbf{x}_i, \mathbf{y}_i) \in \mathcal{D}} (\mathbf{y}_i - \mathcal{N}_{\boldsymbol{\rho}}(\mathbf{x}_i))^2$$

Step 3 : Adjust the parameters to minimize the loss via (stochastic) gradient descent:

$$\rho_i^{\text{new}} = \rho_i^{\text{old}} - \alpha^{\text{step}} \frac{\partial \mathsf{l}_{\boldsymbol{\rho}^{\text{old}}}}{\partial \rho_i^{\text{old}}}$$

the step-width α^{step} is called the "learning rate" and the process of adjusting the parameters is called "learning".

Our loss function

As a loss function, we implement

$$\mathsf{l}_{\boldsymbol{\rho}} := \frac{1}{N_{\text{path length}}} \sum_{\mathbf{x}_i \text{ on sim. path}} \left(\mathbf{G}(\mathbf{x}_i, \mathcal{N}_{\boldsymbol{\rho}})\right)^2$$

where we use $\mathcal{N}_{\boldsymbol{\rho}}$ to simulate a path. $\mathbf{G}$ is chosen, such that the true equilibrium policy $\mathbf{f}(\mathbf{x})$ is defined by $\mathbf{G}(\mathbf{x}, \mathbf{f}) = 0 \; \forall \mathbf{x}$. Therefore, there is no need for labels to evaluate our loss function.

Training deep equilibrium nets[2]

1. Simulate a sequence of states $\mathcal{D}^i_{\text{train}} \leftarrow \{\mathbf{x}^i_1, \mathbf{x}^i_2, \ldots, \mathbf{x}^i_T\}$ from the policy encoded by the network parameters $\boldsymbol{\rho}^i$.
2. Evaluate the errors of the equilibrium conditions on the newly generated set $\mathcal{D}_{\text{train}}$.
3. If the error statistics are not low enough:
   3.1 Update the parameters of the neural network with a gradient descent step (or a variant):

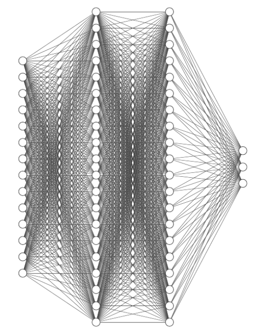
$$\rho^{i+1}_k = \rho^i_k - \alpha_{\text{learn}} \frac{\partial \ell_{\mathcal{D}^i_{\text{train}}}(\boldsymbol{\rho}^i)}{\partial \rho^i_k}.$$

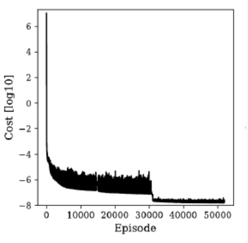
   3.2 Set new starting states for simulation: $\mathbf{x}^{i+1}_0 = \mathbf{x}^i_T$.
   3.3 Increase i by one and go back to step 1.

[2]Sample codes here: https://github.com/sischei/DeepEquilibriumNets.

Mapping the Model to DEQN

DEQN architecture: 2 hidden layers; 1024 nodes; selu activation function; minibatch size: 128; Adam optimizer; learning rate $\alpha_{learn} = 10^{-5}$.





$\rightarrow$ Compute first-order conditions and feed them into the DEQN.

Global Uncertainty Quantification

Why Global Uncertainty Quantification?

- **How are quantitative results of IAMs sensitive to a specific parameterization**?
- **Importance ranking** informs the researcher on which parts to focus on when calibrating or extending a model, or the policymaker on which parameters need further scrutiny.
- Today, often "one-at-a-time" approaches tend to be unstructured and suffer from the fact that they are only **local**, that is, highly dependent on the chosen parameter values.

$\rightarrow$ Need for principled global sensitivity analysis (see e.g., Sudret (2008) and Harenberg et al. (2019)).

Global Uncertainty Quantification: 3 Quantities

- We aim at determining which input parameters (or combinations) contribute the most to the uncertainty of the quantities of interest (QoI) such as the SCC.
- Measures for UQ that we use:
  - **Sobol' index** gives prioritization of the uncertain parameters based on the outcome variance explained.
  - **Shapley value** identifies how much model variance can be attributed to the uncertainty in input parameters.
  - **Univariate effect** shows how each parameter affects the outcomes.

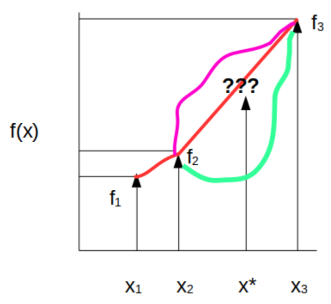
$\rightarrow$ Want to identify which uncertain parameters potentially impact our QoIs based on UQ.

$\rightarrow$ Construct a cheap-to-evaluate **surrogate model** based on Gaussian Processes for QoI's (Scheidegger and Bilionis, 2019), e.g., SCC($\theta_1, \theta_2, ...\theta_N$).

Why Surrogate Models?

- We solved the IAM as a function of exogenous and endogenous states as well as parameters.
- Our computed policies are functions of this extended state space.
- To obtain QoIs as a function of the parameters, we need to simulate the economy with the derived policy functions to compute SCC at N points.
- Solution: Surrogates – high-precision approximations of structural models.

Intuition for GPs



We assume that f's (heights) are Gaussian distributed with zero mean and some covariance matrix K:

$$\begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{bmatrix}\right)$$

Note: f_1 and f_2 should be more correlated due to proximity (compared to f_1 and f_3). Example:

$$\sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 1 & 0.7 & 0.2 \\ 0.7 & 1 & 0.6 \\ 0.2 & 0.6 & 1 \end{bmatrix}\right)$$

Covariance matrix based on a **kernel function**, e.g.,:

$$\kappa(x, x') = \sigma_f^2 \exp\left(-\frac{1}{2\ell^2}(x - x')^2\right)$$

Defining the Gaussian Process Prior

- A GP is defined by a mean function $m(x)$ and a covariance function $k(x, x')$:

$$f(x) \sim \mathcal{GP}(m(x), k(x, x'))$$

- Commonly, we set $m(x) = 0$ to simplify computation, focusing on the covariance function.
- For any set of inputs $\{x_1, \ldots, x_N\}$, the function values $\{f(x_1), \ldots, f(x_N)\}$ follow a multivariate Gaussian distribution:

$$\mathbf{f} \sim \mathcal{N}(0, K)$$

  where $K_{ij} = k(x_i, x_j)$.

Posterior Inference for Gaussian Process Regression

- Given training data $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N$, we want to predict $f(x^*)$ at a new point x^*.
- The joint distribution of observed values $\mathbf{y}$ and the prediction $f(x^*)$ is:

$$\begin{bmatrix} \mathbf{y} \\ f(x^*) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K + \sigma^2 I & k_* \\ k_*^\top & k(x^*, x^*) \end{bmatrix}\right)$$

where $k_* = [k(x^*, x_1), \dots, k(x^*, x_N)]$.

- The conditional distribution of $f(x^*) \mid \mathbf{y}$ is:

$$f(x^*) \mid \mathbf{y} \sim \mathcal{N}(\mu_{f(x^*)}, \sigma^2_{f(x^*)}),$$

where the predictive mean is given by:

$$\mu_{f(x^*)} = k_*^\top (K + \sigma^2 I)^{-1} \mathbf{y}$$

and the predictive variance reads as:

$$\sigma^2_{f(x^*)} = k(x^*, x^*) - k_*^\top (K + \sigma^2 I)^{-1} k_*$$

Predictive Mean and Variance

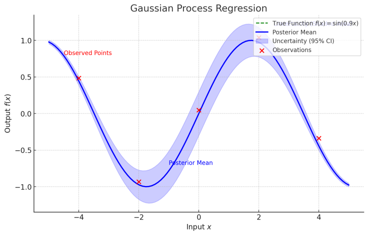
- To interpolate the function at any new point x^*, we use the **predictive mean** of the GP as our best estimate:

$$\text{Predictive Mean:} \quad \mathbb{E}[f(x^*)] = k_*^\top (K + \sigma^2 I)^{-1} \mathbf{y}.$$

- The **predictive variance** quantifies the uncertainty around this estimate, providing confidence intervals:

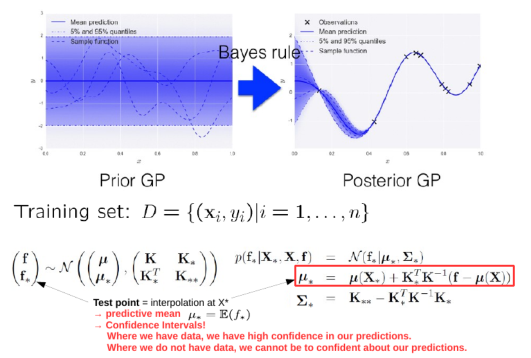
$$\text{Predictive Variance:} \quad \mathrm{Var}(f(x^*)) = k(x^*, x^*) - k_*^\top (K + \sigma^2 I)^{-1} k_*.$$

- **Together, these expressions allow us to both estimate the function's value at any x^* and understand the associated uncertainty**.

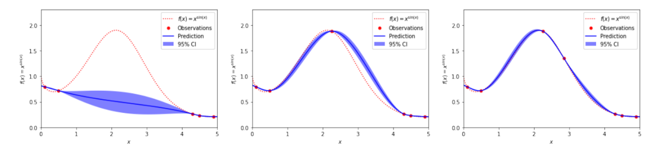


Gaussian Processes in a Nutshell

→ Create a surrogate model for QoIs such as SCC($\theta_1, \theta_2, ...\theta_N$) with GPs.



Bayesian Active Learning



- Training of standard GPs scales as $\mathcal{O}(N^3)$ with the number of observations N.
- Runtime consequently would increase drastically with increasing N.
- Bayesian active learning: "add points where needed the most".

$$U(\tilde{x}) = \sigma_m \tilde{\mu}(\tilde{x}) + \frac{\sigma_v}{2} \log\left(\tilde{\sigma}(\tilde{x})\right) \tag{3}$$

$\sigma_m > 0$, $\sigma_v > 0$, and where $\tilde{\mu}$ and $\tilde{\sigma}$ are the predictive mean and variance of a GP, trained at input locations $\mathbf{X} = \{x_1, \ldots, x_n\}$, and evaluated at $\tilde{x}$, respectively. $\rightarrow$Fit a GP over the $SCC(\theta_1, \ldots, \theta_N)$.

Sobol' indices

- Consider a generic mathematical model $\mathcal{M}(\cdot)$ which has N number of uncertain parameters θ as an input and Q number of model outputs, y summarizes the QoI (e.g., SCC):

$$\theta \in \mathcal{D}_\theta \subset \mathbb{R}^N \mapsto y = \mathcal{M}(\theta_1, \theta_2, \cdots, \theta_N) \in \mathbb{R}^Q.$$

- The first-order Sobol indices are defined as:[3]

$$S_i \equiv \frac{\mathrm{Var}_{\theta_i}\left[\mathbb{E}_{\Theta\setminus\theta_i}[y \mid \theta_i]\right]}{\mathrm{Var}_\Theta[y]}$$

- The nominator $\mathrm{Var}_{\theta_i}\left[\mathbb{E}_{\Theta\setminus\theta_i}[y \mid \theta_i]\right]$ tells you how much the first order effect of θ_i on model output y.
- We normalize the index by the total model variance $\mathrm{Var}_\Theta[y]$ to be scaled in $[0, 1]$.

[3] For simplicity, we assume $Q = 1$.

When to use Sobol' indices?

- Screening: the first-order Sobol' index indicates by what percentage the total variance D would be reduced, should the parameter Θ_i be perfectly known and set to a fixed value.
- If first-order Sobol' index (in practice $< 1\%$): parameter Θ_i could be set to a deterministic value without changing the distribution of the quantity of interest.

Univariate Effect (UE)

- UE: Measures the non-linear relationship between the target QoI and an uncertain input parameter.
- UE: is the conditional expectation, which integrates over the other uncertain parameters θ_{-i}, of QoI as a function of an input parameter θ_i fixed at an arbitrary value ϑ_i:

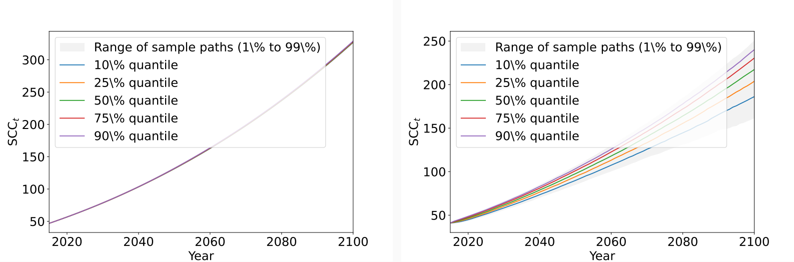
$$\mathcal{M}_i^1(\vartheta_i) = \mathbb{E}_{\theta_{-i}}\left[Y \mid \theta_i = \vartheta_i\right].$$

$\rightarrow$Sobol' indices do not include information about the direction in which it affects the quantities of interest.

$\rightarrow$Univariate effect: They help to find regions of high and low sensitivity, and can be interpreted as a robust direction of change under parameter uncertainty.

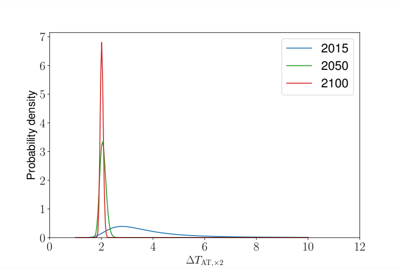
Some Tentative Results

Social Cost of Carbon [$/tonC] - Learning Matters

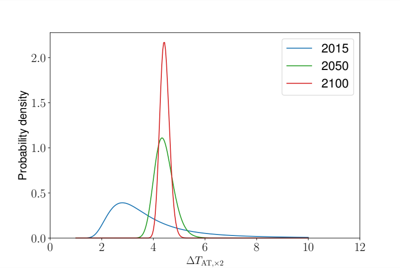


Model with no learning (left) and learning (right); SCC behaves like a random variable, as Temperature and ECS is stochastic (our results confirm Cai and Lontzek (2019)).

Tail learning



(a) $\Delta T^{*}_{\mathrm{AT},\times 2} = 2.0$

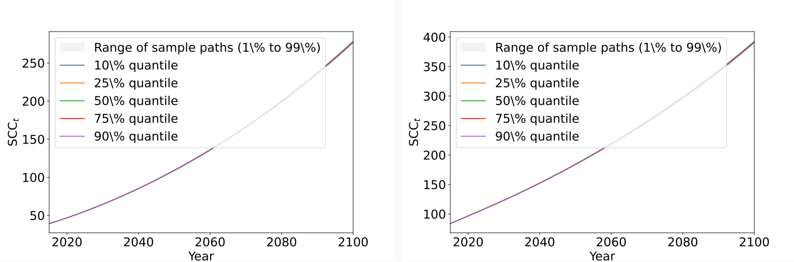


(b) $\Delta T^{*}_{\mathrm{AT},\times 2} = 4.5$

Figure 3: The posterior probability density function of $\Delta T_{\mathrm{AT},\times 2}$ is shown when the true equilibrium sensitivity $\Delta T^{*}_{\mathrm{AT},\times 2}$ is set to 2.0, 3.42, and 4.5, respectively.

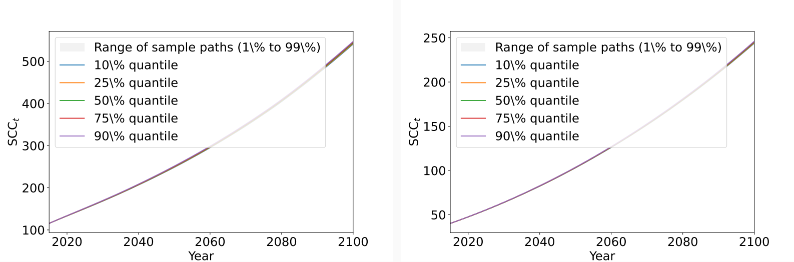
- Social planner learns the upper tail of the prior distribution very quickly (less than a decade for 99% percentile; ECS < 3.42).
- In the case where the ECS is set to 4.5, that is, relatively high, learning slows as the Bayes rule requires more observations to move the mean estimate from the prior ECS to the true high value.

Sensitivity analysis with respect to IES in no learning case



Model with $\psi = 1.05$ (left) and $\psi = 2.0$ (right); (our results confirm Jensen and Traeger (2014)).

Sensitivity analysis with respect to time preference in no learning case

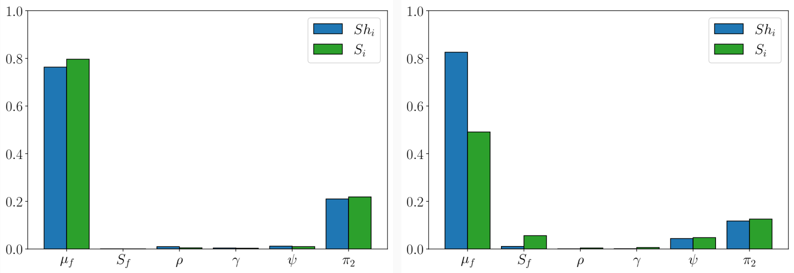


Model with $\rho = 0.01$ (left) and $\rho = 0.02$ (right).

Parameter Ranges for the UQ Analysis

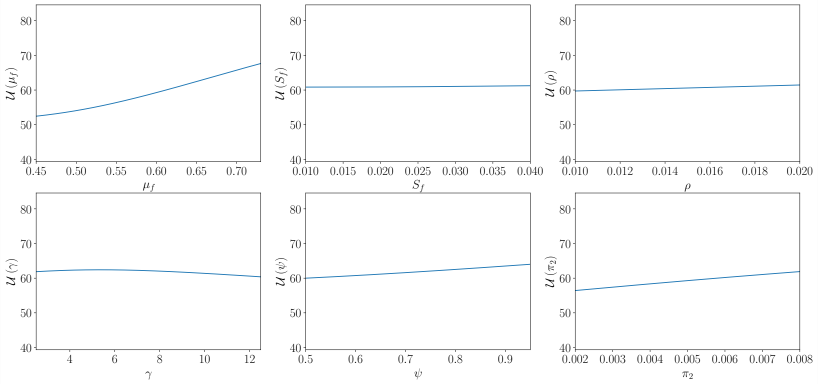
Parameter	θ_i^0	$\underline{\theta}_i$	$\overline{\theta}_i$	Source etc.
ρ	0.015	0.01	0.02	Stern (2007)
γ	10.	5.0	10.0	Jensen and Traeger (2014) and Cai and Lontzek (2019)
ψ	1.5	0.5	2.0	–//–
π_2	0.00236	0.002	0.008	Nordhaus (2017b) and Weitzman (2012)
$\mu_{f,0}$	0.65	0.45	0.73	Roe and Baker (2007)
$S_{f,0}$	0.0169	0.01	0.04	Roe and Baker (2007)

Importance of Parameters for SCC in 2020



UQ results for the model without learning (left) and with learning (right) in 2020.

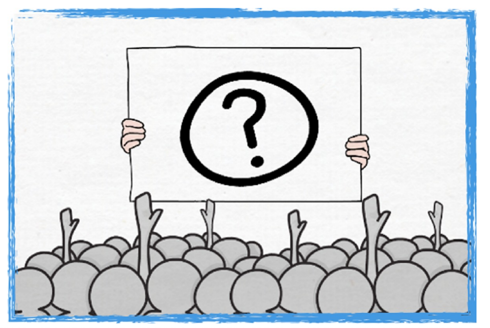
Learning: Univariate effects on the SCC in 2020



Conclusion & Outlook

- Uncertainty in SCC is mainly driven by parametric uncertainty in the ECS and the damage function.
- Overall, highly nuanced results.
- We provide, to the best of our knowledge, the most comprehensive and scalable computational framework to solve large-scale dynamic stochastic integrated models and perform UQ.
- **Some open-source codes here**:
  - `https://github.com/ClimateChangeEcon/Climate_in_Climate_Economics`
  - `https://github.com/sischei/DeepEquilibriumNets`
- Moving forward: We intend to add multiple global regions to the model to study the distributional effects of climate change.

Questions?



References

Ackerman, F., Stanton, E. A., and Bueno, R. (2010). Fat tails, exponents, extreme uncertainty: Simulating catastrophe in dice. *Ecological Economics*, 69(8):1657–1665.

Anderson, B., Borgonovo, E., Galeotti, M., and Roson, R. (2014). Uncertainty in Climate Change Modeling: Can Global Sensitivity Analysis Be of Help? *Risk Analysis*, 34(2):271–293.

Anthoff, D., Tol, R. S., and Yohe, G. W. (2009). Discounting for climate change. *Economics*, 3(1):20090024.

Azinovic, M., Gaegauf, L., and Scheidegger, S. (2022). Deep equilibrium nets. *International Economic Review*, n/a(n/a).

Butler, M. P., Reed, P. M., Fisher-Vanden, K., Keller, K., and Wagener, T. (2014). Identifying parametric controls and dependencies in integrated assessment models using global sensitivity analysis. *Environmental Modelling & Software*, 59:10–29.

Cai, Y. and Lontzek, T. (2019). The social cost of carbon with economic and climate risks. *Journal of Political Economy*.

Folini, D., Friedl, A., Kübler, F., and Scheidegger, S. (2024). The Climate in Climate Economics*. *The Review of Economic Studies*, page rdae011.

Gillingham, K. and Stock, J. H. (2018). The cost of reducing greenhouse gas emissions. *Journal of Economic Perspectives*, 32(4):53–72.

Harenberg, D., Marelli, S., Sudret, B., and Winschel, V. (2019). Uncertainty quantification and global sensitivity analysis for economic models. *Quantitative Economics*, 10(1):1–41.

Jensen, S. and Traeger, C. (2014). Optimal climate change mitigation under long-term growth uncertainty: Stochastic integrated assessment and analytic findings. *European Economic Review*, 69(C):104–125.

Lemoine, D. and Traeger, C. P. (2016). Ambiguous tipping points. *Journal of Economic Behavior & Organization*, 132(PB):5–18.

Miftakhova, A. (2021). Global sensitivity analysis for optimal climate policies: Finding what truly matters. *Economic Modelling*, 105:105653.

Nordhaus, W. D. (2017a). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences of the United States of America*, 114(7):1518–1523.

Nordhaus, W. D. (2017b). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, page 201609244.

Roe, G. and Baker, M. (2011). Comment on" another look at climate sensitivity" by zaliapin and ghil (2010). *Nonlinear Processes in Geophysics*, 18(1):125–127.

Roe, G. H. and Baker, M. B. (2007). Why is climate sensitivity so unpredictable? *Science*, 318(5850):629–632.

Scheidegger, S. and Bilionis, I. (2019). Machine learning for high-dimensional dynamic stochastic economies. *Journal of Computational Science*, 33:68 – 82.

Stern, N. H. (2007). *The economics of climate change: the Stern review*. Cambridge University Press.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety*, 93(7):964–979.

Weitzman, M. L. (2012). Ghg targets as insurance against catastrophic climate damages. *Journal of Public Economic Theory*, 14(2):221–244.

References v

Zaliapin, I. and Ghil, M. (2010). Another look at climate sensitivity. *Nonlinear Processes in Geophysics*, 17(2):113–122.

Zaliapin, I. and Ghil, M. (2011). Reply to roe and baker's comment on" another look at climate sensitivity" by zaliapin and ghil (2010). *Nonlinear Processes in Geophysics*, 18(1):129–131.